A consistent theory for linear waves of the Shallow Water Equations on a rotating plane in mid-latitudes

Nathan Paldor[1, 2], Shira Rubin[1] and Arthur J. Mariano[2]

[1] Institute of Earth Sciences, Hebrew University of Jerusalem; Jerusalem, 91904 ISRAEL

[2] Division of Meteorology and Physical Oceanography; Rosensteil School of Marine and Atmospheric Sciences, University of Miami; 4600 Rickenbacker Causeway Miami, FL 33149 USA

17 January 2006
JPO– In Submission

Abstract

The present study provides a unified and consistent theory for the three types of linear waves of the Shallow Water Equations (SWE) on the β-plane – Kelvin, inertia-gravity (Poincare) and planetary (Rossby). The unified theory obtains by formulating the linearized SWE as an eigenvalue problem that is a variant of the classical Schrödinger equation. The results of the new theory show that Kelvin waves exist on the β-plane with vanishing meridional velocity, as is the case on the f-plane, without any change in the dispersion relation while the meridional structure of their height amplitude is trivially modified from Exponential on the f-plane to a1-sided Gaussian on the β-plane. Similarly, inertia-gravity waves are only slightly modified in the new theory compared to their characteristics on the f-plane. However, for planetary (Rossby) waves (that exist only on the β-plane) the new theory yields similar dispersion relation to the classical theory only for large values of gravity waves' speed. On the other hand, for low gravity-wave phase speed, i.e. in equivalent-barotropic cases with small density jump at the interface, the dispersion relation of the new theory has phase speeds that are twice larger than in the classical theory for realistic widths and up to 3.3 times larger for wide channels. This faster phase propagation is consistent with recent observation of the westward propagation of crests and troughs of Sea Surface Height made by the altimeter aboard the Topex/Poseidon satellite. The unified theory also admits inertial waves, i.e. waves that oscillate at the local inertial frequency, as a consistent solution of the eigenvalue problem.

I. Introduction

The Shallow Water Equations (hereafter, SWE) provide the very fundamental description of the dynamics of an incompressible fluid that occupies a sufficiently thin layer such that the horizontal velocity is uniform across the layer's height. Mathematically, the SWE are nothing but Euler equations for a compressible gas in which the pressure is quadratic with the density but with the density of the gas replaced by the fluid height. Linear waves of the SWE in the presence of rotation fall traditionally into two categories: The first is high-frequency waves (Kelvin waves and inertia-gravity, or Poincaré, waves) that represent rotationally modified gravity waves of the non-rotating SWE. The second type is the low-frequency, planetary (Rossby), waves that originate from the dependence of the Coriolis frequency on latitude: f(y). The derivation of the former type in the context of the SWE is done straightforwardly on the f-plane, where f(y) is replaced by a constant f_0. In contrast, Rossby waves are derived on the β-plane by making additional simplifying assumptions on the flow, e.g., near non divergence or quasi-geostrophy, both of which are consistent with the smallness of $f(y)-f_0 \equiv \beta y$ compared to f_0 (Pedlosky, 1979; Gill, 1982; Cushman-Roisin, 1994).

Although the dispersion relation of planetary (Rossby) waves can be derived directly from the linearized form of the conservation of potential vorticity the meridional structure of the velocity and height eigenfunctions are derived only from a perturbation expansion near a simple (steady, geosptrophic) state. The small expansion parameter used for evaluating the eigenfunctions is β but both the deviation of the velocity from geostrophy and the velocity field's divergence are assumed small to the same order. Both the dispersion relation of Rossby waves and the heuristic explanation for their westward propagation are based on vorticity conservation so changes in the relative vorticity are essential for their existence (see Fig. 3.16.1 in Pedlosky, 1979). However, the flow divergence in these waves is an essential physical element without which the velocity field is time-independent – geostrophic (see Sec. 12.3 and Fig. 12.2 in Gill, 1982). Thus, while Kelvin and inertia-gravity, waves are derived

directly from the SWE without making any assumption on the form of the solutions, Rossby (planetary) waves can only be derived by making some assumptions on the solutions. These assumptions limit the generality of the solutions and imply that each type of waves originates from a different physical set-up, that translates into a different set of mathematical equations.

In the present study we provide a canonical theory that yields the three types of waves: Kelvin, inertia-gravity (Poincaré) and planetary (Rossby), straightforwardly from the SWE without making any additional assumption. This theory yields the following theoretical advances: I) A derivation of the Kelvin and Poincaré wave solutions on the β-plane. II) A derivation of Rossby wave solution that includes the variation of f(y) everywhere and does not let $f=f_0$ in some terms while $\beta\neq 0$. III) Faster phase speed of Rossby waves in the observationally relevant range of parameter values. The last result is in accordance with the altimeter observations made aboard the Topex/Poseidon satellite, which show that Rossby waves in the thermocline of the ocean propagate westward faster than predicted by the classical theory (Chelton and Schlax, 1996; Oschiny and Cornillon, 2004).

2. Linear waves of the Shallow Water Equations and the eigenvalue equation

In vectorial form the linearized SWE with rotation are given by:

$$\begin{aligned} &\frac{\partial \underline{V}}{\partial t} + f\hat{k}\times\underline{V} = -g\nabla\eta, \\ &\frac{\partial \eta}{\partial t} = -H\nabla\cdot\underline{V}. \end{aligned} \tag{2.1}$$

Here, f is the Coriolis frequency, H the unperturbed height (thickness) of the shallow layer of fluid and η is the deviation of height, h, from H (i.e. $h=H+\eta$); ∇ is the two dimensional nabla operator; $\underline{V}$ is the two-dimensional (horizontal) velocity vector; t is time and $\hat{k}$ is a unit vector in the direction perpendicular to the velocity vector $\underline{V}$; g is the gravitational constant in barotropic cases and the reduced gravity (i.e. $g'=g\Delta\rho/\rho_0$ where $\Delta\rho/\rho_0$ is the relative density difference between the lower and upper layers) in equivalent-barotropic cases.

In Cartesian (x, y) coordinates, where x (Rspct. y) is directed eastwards (Rspct. northwards), for $\underline{V}$ =(u, v), where u (Rspct. v) is the velocity components in the eastward (Rspct. northward) direction and for linearly varying Coriolis parameter $f(y) = f_0+\beta y = 2\Omega(\sin(\phi_0) + [\cos(\phi_0)/R]y)$ (where ϕ_0 is a mean latitude and Ω and R are Earth's rotation frequency and radius, respectively) the scalar form of these vectorial equations is:

$$\begin{aligned}
&\frac{\partial u}{\partial t}-(f_0+\beta y)v=-g\frac{\partial \eta}{\partial x},\\
&\frac{\partial v}{\partial t}+(f_0+\beta y)u=-g\frac{\partial \eta}{\partial y},\\
&\frac{\partial \eta}{\partial t}=-H\left(\frac{\partial u}{\partial x}+\frac{\partial v}{\partial y}\right).
\end{aligned} \tag{2.2}$$

Since Earth's radius is the only length scale in the equations we take it to be the length scale in nondimensionalizing these equations. The time scale is $(2\Omega)^{-1}$ and these length and time scales imply the velocity scale $2\Omega R$. If, in addition, we scale the height, h and η=h-H by the mean height H then the nondimensional equations corresponding to system (2.2) are:

$$\begin{aligned}
&\frac{\partial u}{\partial t}-\left(\sin\left(\phi_0\right)+\cos\left(\phi_0\right)y\right)v=-\alpha\frac{\partial \eta}{\partial x},\\
&\frac{\partial v}{\partial t}+\left(\sin\left(\phi_0\right)+\cos\left(\phi_0\right)y\right)u=-\alpha\frac{\partial \eta}{\partial y},\\
&\frac{\partial \eta}{\partial t}=-\left(\frac{\partial u}{\partial x}+\frac{\partial v}{\partial y}\right),
\end{aligned} \tag{2.3}$$

where $\alpha=gH/(2\Omega R)^2$ is the only parameter of the nondimensional model that augments the four dimensional parameters: g, H, Ω and R. The reader is reminded that although the variables in systems (2.2) and (2.3) are designated by the same symbols they are dimensional in the former and nondimensional in the latter and that the nondimensional y coordinate measures the latitudinal distance from ϕ_0 in radians. For the rest of this work nondimensional parameters and variables will be used unless otherwise explicitly stated. System (2.3) applies to flows with length scales, L, satisfying H<<L<R the former condition is a general validity condition of the SWE and the latter originates from the β-plane assumption where higher

order terms in the Coriolis frequency are neglected. The time scale, T, for the validity of the system requires that T>>1/N where N is the frequency of oscillation due to the stratification i.e. Brunt-Vaisala frequency in a continuously stratified ocean and $(g'/H')^{1/2}$ in a 2-layer ocean.

Anticipating linear wave solutions of system (2.3) we let that u, v and η vary with x and t as a zonally propagating wave with wavenumber k and phase speed C i.e. $e^{ik(x-Ct)}$. For this form of (x, t) dependence in system (2.3), the u-equation yields u as a linear combination of V(y)=*i*v(y)/k and η(y):

$$u = \frac{\sin(\phi_0) + \cos(\phi_0) y}{C} V + \frac{\alpha}{C} \eta. \tag{2.4}$$

Substituting this expression for u in the latter two equations of system (2.3) and rearranging the terms one gets the following two linear 1st-order ordinary differential equations:

$$\frac{dV}{dy} = \frac{\sin(\phi_0) + \cos(\phi_0) y}{C} V + (\frac{\alpha}{C} - C)\eta, \tag{2.5a}$$

$$\frac{d\eta}{dy} = \frac{k^2 C^2 - \left(\sin(\phi_0) + \cos(\phi_0) y\right)^2}{\alpha C} V - \frac{\sin(\phi_0) + \cos(\phi_0) y}{C} \eta. \tag{2.5b}$$

Up to this point no assumption was made on the nature of the solutions, e.g. quasi-geostrophy near non-divergence, or the smallness of βy compared to f_0 (the dimensional parameters f_0 and βy appear as $\sin(\phi_0)$ and $\cos(\phi_0)y$, respectively in the nondimensional system 2.4-2.5).

We should point out an important difference between the present theory and the classical theory. For sufficiently small C (e.g. Rossby waves) Eq. (2.5a) implies that $|\eta/V| \sim O(f/\alpha) \sim 1$ and Eq. (2.4) then yields $u/V \sim O(1/C) >> 1$. In contrast, in the classical theory the ratio u/V is derived from the assumption of near non-divergence so $u \sim dV/dy \sim lV$ (where *l* is the meridional wavenumber) i.e. u/V is order 1 instead of order $C^{-1} >> 1$.

In order to solve system (2.5) one needs to specify appropriate boundary conditions. The most natural such conditions are that V(y) should vanish along two values of y, which implies that two zonal walls restrict the y-domain. Thus, the differential system (2.5) and the

associated boundary conditions constitute an eigenvalue problem for small amplitude waves that develop in a channel on the β-plane. The channel is centered at latitude ϕ_0, which defines the mean Coriolis frequency $\sin(\phi_0)$ (dimensionally $2\Omega\sin(\phi_0)$), and the channel walls are located at latitudes $\phi_{walls}= \phi_0 \pm \delta\phi$ so the boundary conditions are $V(y=\pm\delta\phi)=0$. The Coriolis frequency, $\sin(\phi_0)+\cos(\phi_0)y$ (i.e. the nondimensional f(y)), varies linearly with y from a minimal value of $\sin(\phi_0)-\cos(\phi_0)\delta\phi$ in the south wall to a maximal value $\sin(\phi_0)+\cos(\phi_0)\delta\phi$ in the north wall. While system (2.5) can be formally applied to an infinite y-domain, the neglect of higher order terms in the expansion of $\sin(\phi-\phi_0)$ and the neglect of the metric terms of the spherical Earth can not be justified for large values of y.

One immediately notices that when $C^2=\alpha$ in system (2.5) the coefficient of the η term in the V-equation, (2.5a), vanishes so V(y) solves independently of η(y) in Eq. (2.5b). Since the solution of the 1st order equation for V(y) is exponential (see the solution for η(y) below) the vanishing of V(y) at any point, y, (e.g. one of the channel walls) implies that V(y) has to vanish identically (a meridional velocity, V(y), of one sign is not a physically acceptable solution). The substitution V(y)=0 in Eq. (2.5b) yields a similar 1st order equation for η(y) (the equations differs only by the sign of C) that solves exactly into:

$$\eta(y)=\eta_0 e^{\frac{-\left[\sin(\phi_0)+\cos(\phi_0)y\right]^2}{2C\cos(\phi_0)}}. \qquad (2.6)$$

Thus, for $C=+(gH)^{1/2}$ the corresponding height amplitude, η(y), decreases monotonically with y, i.e. the maximal height amplitude of this eastward propagating wave is located on the channel's southern wall. In contrast, for $C=-(gH)^{1/2}$ the height, η(y), increases monotonically with y, i.e. the maximal height amplitude of this eastward propagating wave is located on the channel's north wall. These are the, well-known, pair of Kelvin waves whose phase speed is that of non-rotating gravity waves, while rotation only determines the variation of the height profile with y: η(y). Eqs. (2.4) with V=0 and (2.6) imply that although the Coriolis frequency

is y-dependent in both positive and negative modes the zonal velocity, u(y), is in geostrophic balance with the slope of the height amplitude, dη/dy:

$$u(y)=\frac{\alpha}{C}\eta=-\frac{\alpha}{\sin(\phi_0)+\cos(\phi_0)y}\frac{\partial\eta}{\partial y}=-\frac{\alpha}{f(y)}\frac{\partial\eta}{\partial y}.$$

Another degenerate case is the steady state of system (2.3) which is easily studied directly from Eq. (2.3) by setting to zero all time derivatives there. However, this case is not lost in the transformation to Eq. (2.4, 2.5a, b) and multiplying each of these equations by C and setting C=0 yields V=-(α/f)η (with f=f(y)) in all 3 equations. This relation of V and η is nothing but the geostrophic relation, v(y)=(α/f(y))dη/dx, for V=*i*v/k and η=(-*i*/k)dη/dx. The geostrophic relation for u: u(y)=-(α/f(y))dη/dy, results from substituting of Eq. (2.4) in Eq. (2.5b) and setting $k^2C^2=0$. The degeneracy of C=0 is clear when system (2.5) is written as:

$$C\begin{pmatrix}V\\ \eta\end{pmatrix}_y=\begin{pmatrix}f & \alpha-C^2\\ \dfrac{k^2C^2-f^2}{\alpha} & -f\end{pmatrix}\begin{pmatrix}V\\ \eta\end{pmatrix}$$

so the degeneracy is given by the determinant of the matrix on the RHS:

$$C^2\left(-k^2-\frac{f^2}{\alpha}+\frac{k^2C^2}{\alpha}\right),$$

which vanishes for C=0, i.e. the 2-dimensional system (2.5) is degenerate in this case.

For all values of $C\neq\pm(gH)^{1/2}$ and C≠0 system (2.5) is non-degenerate (i.e. its V(y) and η(y) solutions are coupled) so the two, 1st order equations can be transformed to a single, second order equation. Taking the y-derivative of Eq. (2.5a) and employing Eq. (2.5a) to eliminate η(y) and Eq. (2.5b) to eliminate dη/dy from the resulting equation yields:

$$\frac{d^2V}{dy^2}-\left(\frac{\cos(\phi_0)}{C}+k^2(1-\frac{C^2}{\alpha})+\frac{\left(\sin(\phi_0)+\cos(\phi_0)y\right)^2}{\alpha}\right)V=0;\quad V\left(y=\pm\delta\phi\right)=0. \qquad (2.7)$$

A generalization of this equation to continuously stratified oceans for solutions that are not necessarily zonally propagating wave is given in Eq. 5.15 of LeBlond and Mysak (1978).

Due to its complexity the generalized equation is solved for its vertical modes only while the dispersion relation of its horizontal modes is derived only for the case $f(y)=f_0$ (see below).

In classical linear wave theory the (nondimensional) Coriolis frequency in the last term on the Left Hand Side (LHS) of Eq. (2.7), $f(y)=\sin(\phi_0)+\cos(\phi_0)y$, is replaced by its value at the channel's center, $f_0=\sin(\phi_0)$ (i.e. by omitting the $\cos(\phi_0)y$ term from f(y)), in which case no coefficient in this equation is y-dependent. In this (rather artificial) case, the solutions of the constant-coefficient equation satisfying the boundary conditions at $y=\pm\delta\phi$ are given simply by: $V_n(y)=V_0\sin(\pi(n+1)(y+\delta\phi)/(2\delta\phi))$ for n=0, 1, 2,… for an arbitrary V_0. Since for this solution $V_{yy}=-[(n+1)\pi/(2\delta\phi)]^2V$ the phase speeds, C, are given by the roots of the cubic:

$$\frac{\cos(\phi_0)}{C}+\frac{(n+1)^2\pi^2}{4(\delta\phi)^2}+k^2+\frac{\sin^2(\phi_0)}{\alpha}-\frac{k^2C^2}{\alpha}=0. \qquad (2.8)$$

The dispersion relation for the (slow) Rossby waves is obtained from Eq. (2.8) by assuming $C^{-1}>>\alpha>>C^2$, while for the fast, inertia-gravity, (Poincaré) waves the dispersion relation obtains by assuming $C^2>>\alpha>>C^{-1}$. The resulting expressions of $C(k; \alpha, \delta\phi, \phi_0)$ for these two waves are precisely the nondimensional counterpart of the dimensional expressions found in any of the textbooks on the subject: Pedlosky (1979); Gill (1982) and Cushman-Roisin (1994). The dimensional form of Eq. (2.8) is given in Eq. 15.18 LeBlond and Mysak (1978).

The goal of the present study is to extend the aforementioned classical theory to the case where the βy term ($=\cos(\phi_0)y$) is not neglected compared to f_0 ($=\sin(\phi_0)$). In this case the coefficients of Eq. (2.7) are not constant and the solutions, V(y), are not harmonic oscillations across the channel. However, a solution of Eq. (2.7) satisfying the boundary conditions at the channel walls will still yield the dispersion relation $C(k; \alpha, \delta\phi, \phi_0)$ but in a more complex expression than Eq. (2.8).

The first step in solving Eq. (2.7) is to transform its independent variable, y, to $z=y/(\delta\phi)$ so as to introduce the parameter $\delta\phi$ (half the channel width) that appears in the

boundary conditions directly into the differential equation while the boundary conditions are applied at z=±1. After some trivial re-arranging Eq. (2.7) and the corresponding boundary conditions are written in terms of z as:

$$\varepsilon^2 \frac{d^2V}{dz^2} + \left(E - \left(1+bz\right)^2\right)\ V = 0; \qquad \text{V(z=±1)=0,} \qquad (2.9)$$

where the new parameters are:

$$\varepsilon = \frac{\sqrt{\alpha}}{\sin\left(\phi_0\right)\delta\phi}, \qquad (2.10a)$$

$$b = \frac{\cos\left(\phi_0\right)\delta\phi}{\sin\left(\phi_0\right)}, \qquad (2.10b)$$

$$E = -\frac{\alpha}{\sin^2\left(\phi_0\right)}\left(\frac{\cos\left(\phi_0\right)}{C} + k^2\left(1 - \frac{C^2}{\alpha}\right)\right). \qquad (2.10c)$$

The parameter ε in Eq. (2.9) (defined in Eq. 2.10a) is the ratio between the (nondimensional) radius of deformation, $\alpha^{1/2}/\sin(\phi_0)$, and the (nondimensional) channel half-width, δϕ. Thus, the value of ε is unrestricted from a mathematical viewpoint but for typical values of α in the ocean and for wide channels (where the β-effect is important) it should be of order 0.05. The parameter b as defined in Eq. 2.10b is the maximal relative change in f(y) in the channel: $\delta\phi\cos(\phi_0)/\sin(\phi_0)$, which is always less than π/4 (for $\phi_0=\pi/4$). A solution of the Schrödinger equation (2.9) yields the eigensolution, made up of the eigenfunction V(z) and the associated eigenvalue E(ε, b) (defined in Eq. 2.10c but calculated with no reference to the values of C and k) from which the dispersion relation, C(k), is determined by inverting Eq. (2.10c) to get a cubic C(E) relation (see Eq. 3.1 below).

The general solution of the differential equation (2.9) can be expressed as a linear combination of the Parabolic Cylinder Functions (see Chapter 19 in Abramowitz and Stegun, 1972). However, the eigenvalues are determined by applying the boundary conditions V=0 to a linear combination of these functions, which is as complicated as constructing the solution

by numerically integrating the equation. Below, we will solve the linear eigenvalue problem for E, Eq. (2.9), analytically for special values of b and numerically for general b and ε values and deduce from these solutions the sought dispersion relation C(E). Before doing so we draw some qualitative consequences of the unified formulation Eq. (2.9).

3. Qualitative consequences of the unified formulation

Since Eq. (2.9) is a Sturm-Liouville problem with $p(z)=\varepsilon^2$, $r(z)=1$ and $q(z)=-(1+bz)^2$ (see section 1.8 in Bender and Orszag, 1978 for notation and for details of the following brief discussion) Sturm's theorem ensures that it has an infinite number of eigensolutions, (E_n, $V_n(z)$, n=0, 1, 2,…). All its eigenvalues, E_n, are real positive with $E_n \to \infty$ when $n \to \infty$ and the associated eigenfunctions, $V_n(z)$, have exactly n zeros between z=-1 and z=+1 (so $V_0(z)$ has one sign throughout -1< z <1). Each eigenvalue, E, yields three C(E) roots via Eq. 2.10c:

$$\frac{k^2}{\alpha}C^3 - \left(\frac{E\sin^2(\phi_0)}{\alpha} + k^2\right)C - \cos(\phi_0) = 0. \qquad (3.1)$$

This equation determines $C(k; E, \phi_0, \alpha)$ (where $E=E(\varepsilon(\delta\phi, \phi_0, \alpha), b(\delta\phi,\phi_0)) > 0$ is the eigenvalue of Eq. 2.9) and is the counterpart of Eq. (2.8) of the classical, b=0, theory. The main difference between Eq. (2.8) and (3.1) is that the former results from the application of the boundary conditions to the analytic solution of the differential equation (2.7) when the β term is neglected, while the latter results from the general features of solutions of the exact eigenvalue problem (2.9) without solving it (as it can only be solved numerically; see Sec. 4).

The dispersion relation for the slow (i.e. Rossby, Planetary) waves is given by the small C root of Eq. (3.1). An approximate value for this small-C root obtains by neglecting the k^2C^3/α term in Eq. (3.1) compared to the k^2C term there (recall that $C^2 \ll \alpha$ for Rossby waves). Solving for C one gets:

$$C_n^{Rossby} \approx \frac{-\cos(\phi_0)}{k^2 + \frac{\sin^2(\phi_0)}{\alpha}E_n}. \qquad (3.2)$$

From this expression for the phase speed and from the fact that E_n is an increasing series with n, it is clear that the first, n=0, mode has the largest (in absolute value) phase speed.

The dispersion relation for the fast (Poincaré, inertia-gravity) waves, with $C^2>\alpha$, obtains easily from Eq. (3.1) by dropping the $\cos(\phi_0)$ term and dividing the resulting equation through by C ($\neq 0$). One then gets the quadratic equation expression:

$$\left(C_n^{Poincare}\right)^2 \approx \alpha + E_n \frac{\sin^2\left(\phi_0\right)}{k^2}. \tag{3.3}$$

Since the phase speed of the Poincaré waves is larger than $\alpha^{1/2}$ while that of Rossby waves is smaller (in absolute value) than $\alpha^{1/2}$ (these speeds are separated by the phase speed of the westward propagating Kelvin wave, $C=-\alpha^{1/2}$) Eqs. (3.2) and (3.3) provide fairly accurate approximations to the roots of the cubic equation (3.1).

In addition to these two wave types there are two degenerate cases of the unified equation that were already highlighted in Sec. 2 – the steady, C=0, solution and the Kelvin wave solutions, $C^2=\alpha$. Both solutions appear as special cases of Eq. (3.1) when one lets:

$$Ck^2\left(\frac{C^2}{\alpha}-1\right)=0. \tag{3.4}$$

However, this equation also implies, according to Eq. (3.1) that:

$$CE\frac{\sin^2(\phi_0)}{\alpha}=-\cos(\phi_0), \tag{3.5}$$

which can be satisfied for C=0 only on the f-plane (where $\beta=\cos(\phi_0)=0$) and for $C^2=\alpha$ only by the $C=-\alpha^{1/2}$ root (the negative Kelvin solution).

A solution of Eq. (2.9) includes, in addition to the eigenvalues E_n, their associated eigenfunctions, $V_n(z)$, so the eigensolution is independent of C (which is derived from E_n). Therefore, the eigenfunctions $V_n(z)$ are identical in the two waves so for the same meridional wavenumber, n, the $V_n(y)$ of Poincaré waves is precisely that of Rossby wave. On the other hand, the $u_n(z)$ and $\eta_n(z)$ eigenfunctions, are different for the two waves since they are related

to $V_n(z)$ by the phase speed, C, which is different for the two waves (see Eqs. 2.4 and 2.5a). These qualitative consequences are valid even though no assumption was made on the smallness of either the β term or the divergence or the ageosptrophic velocity component.

4. Eigenvalues of equation (2.9) and the corresponding phase speeds

Although Eq. (2.9) is a Schrödinger equation, which has been studied extensively in theoretical physics it has no known solutions for any ε and b. The reason is that Eq. (2.9) applies on the finite interval -1<z<1, where the potential $(1+bz)^2$ has no symmetry (Fig. 1), but at z=±1 the boundary condition is V=0 so an infinite potential well exists at z=±1. In the special cases b=0 and b=1 the potential is symmetric (Fig. 1) so analytic solutions can be found while for 0<b<1 Eq. (2.9) can be easily integrated numerically from z=-1 to z=+1 (no singular points exist) so the eigensolution can be found numerically.

4.1. Analytic solution for b=0

For b=0 the differential equation (2.9) has constant coefficients so its eigenfunctions can be solved exactly and the eigenvalues, E_n, can be determined from these (purely oscillatory) solutions by applying the boundary conditions. The eigensolutions are then:

$$V_n(z)=\frac{2}{\pi(n+1)}\sin\left(\frac{(n+1)\pi}{2}(z+1)\right),\quad E_n=1+\left(\frac{\pi\varepsilon(n+1)}{2}\right)^2,\qquad \text{n=0,1,2…}\quad (4.1)$$

so that $E_n(\varepsilon, b=0)>1$ for all n. Substituting the expression for E_n in Eq. (3.2) yields the known dispersion relation of planetary waves, derived in Rossby (1939):

$$C_n^{Rossby}=\frac{-\cos(\phi_0)}{k^2+E_n\dfrac{\sin^2(\phi_0)}{\alpha}}=\frac{-\cos(\phi_0)}{k^2+\dfrac{(n+1)^2\pi^2}{(2\delta\phi)^2}+\dfrac{\sin^2(\phi_0)}{\alpha}}. \qquad (4.2)$$

Likewise, Eq. (3.3) yields the dispersion relation for inertia-gravity (Poincaré) waves:

$$\left(C_n^{poincare}\right)^2=\alpha+\frac{E_n\sin^2(\phi_0)}{k^2}=\frac{\sin^2(\phi_0)}{k^2}+\frac{\alpha}{k^2}\left(k^2+\frac{(n+1)^2\pi^2}{(2\delta\phi)^2}\right). \qquad (4.3)$$

The frequency associated with this phase speed, k^2C^2, is the Pythagorean sum of the inertial frequency, $\sin^2(\phi_0)$, and the gravitational frequency, $\alpha\kappa^2$, where $\kappa=(k^2+(n+1)^2\pi^2/(2\delta\phi)^2)^{½}$ is the total (meridional and zonal) wavenumber.

4.2. Power series solutions for b>0

For b>0 the solution of Eq. (2.9) can be written in the form $V(z)=\theta(z)e^{\frac{-(1+bz)^2}{2\varepsilon b}}$ where θ(z) is a solution of the differential boundary value problem:

$$\frac{d^2\theta}{dz^2}-2\left(\frac{1+bz}{\varepsilon}\right)\frac{d\theta}{dz}+\left(\frac{E}{\varepsilon^2}-\frac{b}{\varepsilon}\right)\theta=0;\qquad \theta\left(z=\pm1\right)=0. \tag{4.4}$$

Since Eq. (4.4) is regular for all ε>0 it has a regular series expansion. The boundary condition θ(z=-1)=0 suggests the following series expansion (obtained by a change of variables x=1+z):

$$\theta(z)=\sum_{j=1}^{\infty}a_j\left(1+z\right)^j, \tag{4.5}$$

where the series starts at j=1 (and not j=0) to ensure that θ(z=-1)=0. Substituting this series into Eq. (4.4) and equating like powers of (1+z) yields the recursion relation for $\{a_j\}$:

$$\begin{aligned}
&a_1=1; \qquad \text{(this is a trivial normalization condition)}\\
&a_2=\frac{1-b}{\varepsilon};\\
&a_3=\frac{b}{2\varepsilon}-\frac{E}{6\varepsilon^2}+\frac{2\left(1-b\right)^2}{3\varepsilon^2};\\
&.\\
&.\\
&a_{j+2}=\frac{2\left(1-b\right)\left(j+1\right)a_{j+1}+\left(b\left(2j+1\right)-\frac{E}{\varepsilon}\right)a_j}{\varepsilon\left(j+1\right)\left(j+2\right)};\quad j\geq1.
\end{aligned} \tag{4.6}$$

From the series expansion it is clear that when two of its consecutive coefficients vanish the infinite series terminates at some j and becomes a polynomial in (1+z) (which is also a polynomial in z). The series can be employed for calculating the solution for θ(z) and the parameter E can then be varied to find those values at which θ(z=+1)=0, which is as efficient computationally as direct integration of the ODE (be it Eq. 4.4 or Eq. 2.9) with a standard

high accuracy integration algorithm. However, the series expansion is helpful in finding analytic expressions for the eigenvalue problem in the special b=1 case.

4.3. Analytic solutions for b=1

Although for oceanographic application on a sphere b has to be smaller than $\pi/4 \cong 0.79$ from a mathematical viewpoint, b=1 case is a valid special case on the β-plane. Since b=1 is the only other case (besides b=0) where analytic solution of the eigenvalue problem exists we derive this solution so as to substantiates our numerical solution for general b>0. For b=1 the parabolic potential in Eq. (2.9), $(1+1\cdot z)^2$, is symmetric about z=-1 on the $-3 \le z \le 1$ interval. The change of variables $x=(1+z)/2$, which maps the $-3 \le z \le 1$ interval to the $-1 \le x \le 1$ interval, yields a classical symmetric potential about x=0 where the eigenvalues are those of Hermite equation (see table 22.6 in Abramowitz and Stegun, 1972) $E_m=(2m+1)\varepsilon$. Since we are looking for eigenfunctions that vanish at x=0 (i.e. z=-1) only odd eigenfunctions of the symmetric (Hermite) equation are also solutions of the eigenvalue problem (2.9). Therefore, setting m=2n+1 in the eigenvalues of Hermite equation, $E_m=(2m+1)\varepsilon$ yields the eigenvalues of our problem: $E_n=(4n+3)\varepsilon$.

One can show this simple result directly from the power series expansion, Eq. (4.6), by noticing that for b=1 all even-indexed coefficients, $\{a_{2j+2}\}$, vanish (which reflects the symmetry of Eq. 2.9 about z=-1) so the recursion relation for the odd-indexed coefficients is:

$$
\begin{aligned}
& a_1 = 1; \\
& a_3 = \frac{1}{2\varepsilon} - \frac{E}{6\varepsilon^2} = \frac{3\varepsilon - E}{6\varepsilon^2}; \\
& . \\
& . \\
& a_{2j+3} = \frac{(4j+3)\varepsilon - E}{2j(2j+1)\varepsilon^2} a_{2j+1}; \qquad j \ge 0.
\end{aligned}
\tag{4.7}
$$

The eigenvalues, E_n, are determined by requiring that the infinite series in (4.7) becomes a polynomial of degree 2j+1 (j=0, 1, 2, ...) with odd powers of (1+z), which implies for E_n:

$$E_n=(4n+3)\varepsilon, \tag{4.8}$$

when the mode index n is identified with the polynomial index j. In particular, for n=0:

$$E_0=3\varepsilon. \tag{4.9}$$

Substituting the solution for the eigenvalues, Eq. (4.8), into the dispersion relations of the two waves, Eqs. (3.2)-(3.3), yields the following approximate dispersion relations for b=1:

$$C_n^{Rossby}=\frac{-\cos(\phi_0)}{k^2+E_n\dfrac{\sin^2(\phi_0)}{\alpha}}=\frac{-\cos(\phi_0)}{k^2+\dfrac{\sin(\phi_0)(4n+3)}{(\delta\phi)\sqrt{\alpha}}}, \tag{4.10}$$

and

$$\left(C_n^{Poincare}\right)^2=\alpha+E_n\frac{\sin^2(\phi_0)}{k^2}=\alpha+\frac{\sin(\phi_0)(4n+3)\sqrt{\alpha}}{(\delta\phi)k^2}. \tag{4.11}$$

These dispersion relations for b=1 differ markedly from the corresponding relations for b=0, Eqs. (4.2)-(4.3): The b=0 eigenvalues, $E_n(\varepsilon, 0)$, are all larger than 1.0 (Eq. 4.1) while those for b=1, $E_n(\varepsilon, 1)$, are significantly smaller than 1.0 for sufficiently small ε (Eq. 4.8). What is still unclear at this point is whether this decrease of the eigenvalues $E_n(\varepsilon, b)$ with b for fixed ε is monotonic with the increase in b form 0 to 1. This question can be answered by solving the eigenvalue problem (2.9) numerically for general b-values.

4.4. Numerical calculation of the eigenvalues for general b

A standard (shooting) method for solving the eigenvalue problem consists of integrating the differential equation from the south boundary at z=-1 (starting with V=0 and, say, dV/dz=1 there) to the north boundary at z=1 (we used a 5^{th} order Runge-Kutta algorithm with 10^{-10} tolerance) and varying the values of the parameters so as to satisfy the V(z=+1) boundary condition. Except for ε=0 the solution is regular, as can be verified by the expansion in subsection 4.2, so the numerical integration yields a very accurate value of V(z=+1; E, ε, b) (namely, the value of V(z=+1) for given values of the three parameters). For fixed values of ε and b we find (numerically) the roots of the nonlinear equation 0 = F(E) = V(z=+1; E, ε, b). The resulting $E_0(\varepsilon, b)$ contours are shown in Fig. 2, from which it is easy to

verify that the numerical solutions along the b=0 and b=1 ordinates are exactly those given analytically in Eqs. (4.1) (with n=0) and (4.9), respectively. Two points should be now made with regard to the contours. The first point is that for large ε-values the eigenvalue E_0 varies only slightly with b so $E(\varepsilon, b) \approx E(\varepsilon, 0)$ so that C, too, is close to its value in the classical theory. We have verified (results not shown) that for ε-values larger than 0.6 (ε=2, 10 and 25) the E-contours become even closer to horizontal (i.e. similar values of $\delta E=E(\varepsilon,1)-E(\varepsilon,0)$ but for much larger values of E) so the near-independence of E on b for large ε is not limited to the small range of $0.3<\varepsilon<0.6$ shown in Fig. 2. The second point is that at low values of ε the slopes of the E-contours are all positive, which implies that for small fixed ε an increase in b results in a decrease in E. This decrease of E(b) for fixed and small ε is quite drastic: At ε=0.05, for example: $E(b=1)/E(b=0)\approx 0.2$ and $E(b=0.5)/E(b=0)\approx 0.5$!

A somewhat different view of the $E_n(\varepsilon, b)$ relationship discussed up to this point is obtained by plotting $E_n(b; \varepsilon)$, namely by regarding ε as a parameter in the $E_n(b)$ relationship. Figure 3 shows the resulting $E_0(b)$ curves for the indicated values of ε in the interval $0\le b\le 1$. The three panels clearly demonstrate that for small values of $\varepsilon\le 0.6$ (upper panel) the value of E_0 decreases monotonically with b throughout the entire $0\le b\le 1$ interval while for large values of $\varepsilon\ge 0.75$ the value of E_0 increases monotonically with b there. However, in the narrow intermediate region of ε near 0.7 the variation of $E_0(b)$ is not monotonic in the $0\le b\le 1$ range.

4.5. Phase speed of Rossby waves

The results obtained in the preceding subsection for the eigenvalues of Eq. (2.9) have to be translated into estimates for the phase speed of Rossby waves in order for them to be of any significance in physical oceanography. This requires that the 3 (nondimensional) parameters of system (2.9) – E, ε and b – be transformed to the 5 (nondimensional) parameters of system (2.7) – C, k, $\delta\phi$, ϕ_0 and α. To do so, we fix ϕ_0 at some (midlatitude) value so according to Eqs. (2.10a, b) $\delta\phi=\tan(\phi_0)b$ and $\alpha=(\varepsilon b\tan(\phi_0)\sin(\phi_0))^2$. Any pair of

values of ε and b determines $E_0(\varepsilon, b)$ (via Fig. 2), α (via $\alpha=(\varepsilon b\tan(\phi_0)\sin(\phi_0))^2$) and δϕ (via $\delta\phi=\tan(\phi_0)b$) from which the dispersion relations, C(k), are determined by numerically finding the three roots of (3.1) (as was explained in the beginning of Sec. 3).

The two panels in Fig. 4 compare the exact dispersion curves, C(k), of Rossby waves for indicated values of α, δϕ and ϕ_0 based on the same $E_0(\varepsilon, b)$ curves of Fig. 2. The point of these graphs is that the C values in the new theory can be over 3 times larger than those in Rossby's original theory for a wide channel (δϕ=0.6~34.4°) centered at 45° (near k=5 in the upper panel). Even for a realistic "North Pacific" channel that occupies the range of latitudes between 11.5° and 51.5° (i.e. ϕ_0 = 31.5° and δϕ=20°; lower panel) the new theory yields a phase speed that is twice larger than the b=0 theory.

4.6. Inertial waves

A class of waves that exists in rotating fluids but is of lesser relevance to the ocean (or the atmosphere) is inertial (also called gyroscopic) waves. These waves exist even when the pressure gradient force vanishes identically due to the presence of the Coriolis force. Thus, on the f-plane the waves' frequency is the Coriolis frequency, i.e. $k^2C^2=\sin^2(\phi_0)$ in the present study's notation. Since the phase speed is determined by the solutions of the 2nd-order homogenous differential equation it is impossible for these waves to satisfy two boundary conditions, which the reason for designating them as "spurious" in some textbooks (see e.g. Sec. 3.9.iii in Pedlosky, 1979).

Since the pressure gradient force vanishes for g=0=α inertial motion prevails in the present theory for $\alpha\sim(\varepsilon b)^2\to 0$ so setting $k^2C^2=\sin^2(\phi_0)$ in Eq. (3.1) and rearranging yields:

$$E = 1-\alpha\sin^{-2}\left(\phi_0\right)\left(k^2+\cos\left(\phi_0\right)C^{-1}\right), \qquad (4.12)$$

which implies E≤1 for $\alpha\sim(\varepsilon b)^2<<1$. Our analytical solution for b=0 (Eq. 4.1) yields E=1 only at ε=0, for which value the only solution of Eq. (2.9) is V=0 while E<1 is not a solution of the eigenvalue problem. This explains the "spurious" nature of inertial waves in the traditional,

b=0, theory. On the other hand, our numerically calculated contours in Fig. 2 show that E<1 contours exist in the large ($\varepsilon<<1$, $0<b<1$) domain. Thus, inertial waves are solutions of the eigenvalue problem for b>0, a range that is completely inaccessible by the classical theory.

5. Eigenfunctions and the fields' divergence/vorticity

Although the power series expansion in Eq. (4.5) (with the coefficients given in Eq. 4.6) also provides a way for expressing the eigenfunction V(z) for given values of E the summation of a large number of terms is less efficient (from a numerical viewpoint) than a straightforward integration of the ODE in Eq. (2.9). This summation procedure also suffers from the convergence of the series at 1+z=2 (z=1) that is not guaranteed with a fixed number of terms in the series for all values of E, b and ε.

In solving the eigenvalue problem Eq. (2.9) numerically we use the fact that the differential equation in and its associated boundary conditions are homogeneous so the amplitude of the eigenfunctions is undetermined and the normalization of the eigenfunctions is arbitrary. The eigenfunction that corresponds to an eigenvalue, E_n, is found by integrating the differential equation (2.9) from z=-1 with V=0 and $dV/dz\neq 0$ (the value is arbitrary) to z=+1 and the choice of E_n for E guarantees that V(z=+1) = 0. The accuracy of the 5^{th} order Runge-Kutta scheme was determined a-priori to relative accuracy of 10^{-10}. The corresponding height and zonal velocity eigenfunctions (i.e. $\eta(z)$ and u(z)) are then easily obtained by substituting the numerically found solution for V(z) and dV(z)/dz into Eqs. (2.5a) and (2.4), respectively. One should only replace y in these expressions by $z\delta\phi$ and select the value of C that is relevant to the particular wave: planetary (Rossby) or inertia-gravity (Poincaré). As was noted above the Poincaré and Rossby waves share the same V(z) eigenfunctions.

Based on the horizontal shape of the E-contours in Fig. 2 for sufficiently large ε one expects the classical, b=0, theory (Eq. 4.1 and Sec. 4.1) to provide an accurate approximation for the eigenfunctions for sufficiently small b. For Fixed values of k and ϕ_0 (the results shown

below are for k=1.0=ϕ_0) the parameters of the eigenvalue problem: E_0, ε and b determine uniquely the values of C, α and $\delta\phi$ via the relations 2.10. For large ε (ε=2; b=0.1) where E_0=10.871, Eq. (4.1) provides an excellent estimate for the eigenvalue: $E_0=1+\pi^2$= 10.870 so the relative error in E_0 of the b=0 theory is only 10^{-5} and C_{rosbby}=-0.00478, the same value (to 3 significant digits) as in the classical theory.

In accordance with this accurate estimate of the eigenvalue, the solution for the eigenfunctions shown in Fig. 5 demonstrates that the associated V(y) eigenfunction (upper left panel) is exactly the one predicted by the b=0 theory – a pure sinusoidal variation that vanishes at the boundaries. However, even in this case the u(y) and η(y) that vary across the channel in a nearly sinusoidal manner are not identical with their counterparts of the classical b=0 theory. For Rossby waves (upper right panel) η(y) does not vanish at the boundaries (in accordance with Eq. 2.5a) while u(y) is not exactly 90° out of phase relative to V(y) and, as anticipated in Sec. 2 its amplitude is O(1/C)~200 larger than the amplitude of V (instead of $1/\delta\phi$~7). Similar changes occur for u(y) and η(y) of the Poincaré wave (lower left panel) where a close inspection shows that they are not exactly anti-symmetric with y. The quadratic exponent of the Kelvin wave (lower right panel) in the present theory differs from the linear exponent of the classical theory only slightly and most of the difference occurs in the center of the channel due to the $\eta(y=-\delta\phi)$=1 normalization and the exponential decay of η(y) with distance from the wall.

In contrast, for small ε (ε=0.055, b=0.15) where E_0=0.862 (C_{rossby}=-0.000103) the b=0 theory yields (Eq. 4.1) $E_0=1+0.055^2\pi^2/4$=1.0075, i.e. the eigenvalue is in significant error. It is, thus, not surprising that the structure of the associated eigenfunction, V(y), is also far from pure sinusoidal as anticipated by the classical, b=0, theory (upper left panel in Fig. 6). In fact, it is easy to show that the term $E-(1+bZ)^2$ in Eq. (2.9) is negative for $z > (E^{1/2}-1)/b$ and positive for $-1<z<(E^{1/2}-1)/b$. Therefore, for the present values of E (=0.862) and b (=0.15) V should

decay exponentially for z>-0.48 (i.e. y>-0.48*b*tan(1)=-0.11) and oscillate for -1<z <-0.48 (i.e. -b*tan(1)<y<-0.11) so even near-symmetry about y=0 should not be expected. This asymmetry is expected in light of the expansion of $V_n(z)$ in Sec. 4.2 into a symmetric Gaussian about 1+bz=0 times a power series in (1+z), neither of which is symmetric about z=0. The η(y) and u(y) functions of the Rossby (upper right panel) and Poincaré (lower left panel) waves follow from V(y) with the different values of C and α in Eq. (2.4, 2.5a). The exponent of the Kelvin wave (lower right panel) undergoes a much more drastic change across the channel than in Fig. 5 due to the small radius of deformation (α=0.000117 so $\alpha^{1/2}$/sin(1)=0.0129) compared to the, relatively, wide channel 2δϕ=2btan(1)=0.46.

In the classical linear wave theory Rossby waves can be derived from the vorticity equation and their existence relies upon the β-term so they have no counterpart in nonrotating fluids. In order to solve the problem the divergence field is assumed small (but not zero) and this wave is therefore considered non-divergent. To compare our theory that derives from the SWE without imposing any assumption on the fields' divergence with the classical theory we need to calculate the divergence (δ) and vorticity (ζ) associated with the velocity fields that obtain in our theory. The definitions of these variables and the relation v=-*i*kV imply:

$$\delta = \frac{du}{dx} + \frac{dv}{dy} = ik(u - V_y),$$
$$\zeta = \frac{dv}{dx} - \frac{du}{dy} = k^2 V - \frac{du}{dy},$$

which, upon substituting Eqs. (2.4, 2.5a), yields:

$$\frac{\delta}{ik} = -\frac{fC}{\alpha - C^2} V + \frac{C^2}{\alpha - C^2} V_y,$$
$$\zeta = -\left(\frac{\beta}{C} + \frac{f^2}{\alpha - C^2}\right) V + \frac{fC}{\alpha - C^2} V_y, \qquad (5.1)$$

(f=sin(ϕ_0)+cos(ϕ_0)y; β=cos(ϕ_0)). Upon substituting δ/(*i*k) into ζ, PV conservation yields:

$$ikC\zeta - \beta v - f\delta = 0. \qquad (5.2)$$

The divergence and vorticity curves shown in Fig. 7 and Fig. 8 correspond to the parameters and eigensolutions of Fig. 5 and Fig. 6, respectively. It is evident that for both large (Fig. 7) and small (Fig. 8) values of α the divergence field of Rossby waves is nearly everywhere negligible compared to the vorticity field of these waves. This is a result of the low C value of these waves, which according to the continuity equation makes the divergence field small to order C, or equivalently Eq. (5.2) implies that $\delta/\zeta \sim C$. As is evident from Eq. (5.1) $\delta(y)$ is approximated by $Cfv(y)/\alpha + O(C^2)$ for Rossby waves, which supports the numerical results in the upper panels of Figures 7 and 8. Since for n=0 the V(y) does not vanish inside the channel it is clear that near internal points where the vorticity vanishes (for example when it changes sign) the divergence has to be larger than the vorticity so the smallness of δ/ζ is not uniform throughout the entire channel.

For Poincaré waves (the bottom two panels in Figs. 7 and 8), where C is O(1) so δ and ζ are of the same order so these waves can not be considered non-rotational (as are gravity waves in non-rotating systems). The vorticity and divergence of Kelvin waves, where v(y)=0, vary across the channel as u(y) and du/dy, respectively.

6. Concluding remarks and Summary

The consistent formulation for all linear waves on the mid-latitude β-plane suggested in this study was first suggested in the b=0 case and for Rossby waves only by Lindzen (1967). It was also applied to the equatorial β-plane by Erlick and Paldor (2006) with slight modifications of the eigenvalue problem Eq. (2.9). The equation on the equator can be obtained from Eq. (2.9) by multiplying it through with $\sin^2(\phi_0)$ (=0 on the equator) and replacing the $\cos(\phi_0)$ factor in E and b by 1.0. As a result the $\sin(\phi_0)$ factors in E, ε and b disappear and the potential $(1+bz)^2$ becomes $(\delta z)^2$ when the equation is divided through by ε^2 (with $\delta=(\delta\phi)^2/\alpha^{½}$. The entire problem now rests upon solving the 1-parameter Schrödinger eigenvalue equation $E^*(\delta)$ (where E* is the modified eigenvalue). It turns out that the

analysis of the equatorial case is much simpler compared to the present, mid-latitude, case. The interested reader is referred to Erlick and Paldor (2006) for more details on the application of the Schördinger equation approach to the equatorial β-plane.

In the mid-latitude GFD theory both Kelvin and Poincaré (i.e. inertia-gravity) waves are obtained on the f-plane while Rossby waves are derived on the β-plane. Since the former two waves exist even when for constant planetary vorticity, f, they are classified as non-rotational while Rossby waves, that exist only on the β-plane are classified as non-divergent. The present study alters this traditional view by deriving, for the first time, all three types of waves in a consistent unified theory based on the same physical set-up and governing differential equations. The non-divergence of Rossby waves result from the small C there that according to the continuity equation (where the LHS is $-ikC\eta$) $\underline{\nabla}\cdot\underline{V}$ (on its RHS) is order C.

For Kelvin waves, with V(y)=0, are traditionally derived on the f-plane the present study shows that they exist also on the β-plane and that their dispersion relation, $C=\pm(gH)^{1/2}$, is identical to that on the f-plane. This result can be anticipated since f does not appear in the dispersion relation in the f-plane theory. The height eigenfunctions of Kelvin waves, η(y), is trivially modified on the β-plane compared to its counterpart on the f-plane with f_0 (i.e. $\sin(\phi_0)$) replaced by $f_0+\beta y$ (i.e. $\sin(\phi_0)+\cos(\phi_0)y$) in the first order equation (2.5b with V=0). This leads to a Gaussian decay of η(y) with distance from the channel walls – Eq. (2.6) – whereas on the f-plane this decay is exponential with the distance – $\eta(y)\sim e^{-\sin(\phi_0)y/C}$.

inertia-gravity (Poincaré) waves are also classically developed only on the f-plane and the present study both shows that they also exist on the β-plane and that they originate from the exact same equations that as Rossby waves. This new unified derivation of inertia-gravity waves and planetary (Rossby) waves, where both waves originate from the same solution of the linear eigenvalue equation, (2.9), places the rotational and divergence arguments of their physical origin as a consequence of their different dispersion relations rather than its cause. In

the high frequency inertia-gravity waves where the phase speeds are given by the large C roots of Eq. (3.1) the continuity equation yields fast temporal variations of η so the velocity divergence has to be large as well. The opposite holds for the low frequency planetary waves (where C is the smallest root of Eq. 3.1) where the velocity divergence required to balance the temporal changes in η is small i.e. the flow is nearly non-divergent.

Perhaps the most significant result of the new derivation proposed in this study has to do with Rossby waves. Previously, these waves were only derived in perturbation expansion procedures in which $\beta=df/dy$ was treated as a small parameter that is retained in the governing equations (whether the SWE or the vorticity equation) while, at the same time, f(y) is replaced by f_0 (i.e. βy is neglected compared to f_0) everywhere else in these equations. These two conflicting assumptions are made so as to include the β-effect in the equations but, at the same time, leave the coefficients in these equations constant to ensure that they can be solved analytically. In the present theory $f(y)=f_0+\beta y$ is assumed everywhere and both Rossby waves and inertia-gravity waves from different roots of the same eigenvalue equation. The V(y) eigenfunctions in the classical perturbation theory are all pure trigonometric functions that vanish on the boundaries while in the new theory these functions have a much more complex form (e.g. Hermite functions) that changes with the values of the model parameters. The analysis presented in this study focuses on the first eigenvalue, E_0, but applies also to the higher eigenvalues, i.e. $E_n(\varepsilon, b>0) > E_n(\varepsilon, b=0)$ nearly everywhere. However, since C_n^{rossby} decreases with n (see Eq. 3.2) the effect of letting b>0 is strongest for the n=0 mode.

A somewhat surprising result that is evident in the contours of Fig. 2 is that the classical theory is significantly modified in the present theory even at low values of b. Even for b=0.3 (i.e. $\delta\phi=0.17$ at $\phi_0=\pi/6$) and $\varepsilon=0.05$ (radius of deformation $\alpha^{1/2}/\sin(\phi_0)R$ of about 55 km in a channel of half-width $\delta\phi R$ of 1100 km, where the β term is important) the eigenvalue in the present theory is about $E_0=0.7$ while the classical theory (Eq. 4.1) has $E_0>1$. For k<<

10^3 (so $k^2 << \sin^2(\phi_0)E/\alpha$ in Eq. 4.2) the phase speed of Rossby waves is well approximated by $C \sim -\cos(\phi_0)\alpha/(\sin^2(\phi_0)E)$ so the relative error in C equals that in E_0, i.e. both are about 50%

Inertial waves, whose frequency of oscillation equals f_0, appear in the present theory as the regular solutions of the eigenvalue problem in the range $\alpha \sim \varepsilon b \rightarrow 0$ and $E \leq 1$. This is in contrast to the classical theory where their dispersion relation is that of inertia-gravity waves in the limit g=0 but the corresponding eigenfunction can not satisfy the boundary conditions.

The last point we wish to make regards the application of the present theory's results to observation such as the estimate of the westward motion of SSH features by the altimeter aboard the Topex/Poseidon satellite. Based on the structure of the $\eta(\phi)$ (upper-right panel of Fig. 6) one can argue that the phase speeds of the present theory should be compared with those of the classical theory with β evaluated at a latitude closer to the South wall and not at the center of the channel since this is where the SSH signal (i.e. $\eta(\phi)$) is maximal. Even if one changes the pertinent latitude for determining β from the channel center (ϕ_0=0.55 Rad.≈31.5° in the North Pacific case of Fig. 4) all the way to the South wall ($\phi_0 - \delta\phi$=0.2 Rad.≈11.5°) the corresponding change in β, and with it the increase in phase speed, is only cos(0.2)/cos(0.55) = 1.15. This will increase C, as measured by the slope of the dotted curve in the bottom panel of Fig. 4, ($6.15 \cdot 10^{-4}/20 = 3.08 \cdot 10^{-5}$) by 15% to $3.54 \cdot 10^{-5}$, which leaves the value of C in the present theory ($1.35 \cdot 10^{-3}/20 = 6.8 \cdot 10^{-4}$) about 92% larger than that of the classical theory even when the latter is calculated at the South wall.

Acknowledgement: This work was supported by grant No. 579/05 of the Israel Science Foundation to HU. We thank two anonymous reviewers for their helpful comments.

References

Abramowitz, M. and I. A. Stegun, 1972: *Handbook of mathematical functions.* Dover Publications, Inc., NY, USA 1043 pp.

Bender, C. M. and S. A. Orszag, 1978: *Advanced mathematical methods for scientists and engineers.* McGraw-Hill, Inc., 593 pp.

Chelton, D.B. and M.G. Schlax, 1996. Global observations of oceanic Rossby waves. *Science,* **272,** 234-238.

Cushman-Roisin, B., 1994: *Introduction to Geophysical Fluid Dynamics.* Prentice-Hall, Inc., NJ, USA 210 pp.

Erlick, C. and N. Paldor, 2006: Linear waves in a symmetric zonal channel on the equatorial \beta-plane. *Q. J. Roy. Met. Soc.,* Submitted.

Gill, A. E., 1982: *Atmosphere-Ocean Dynamics.* Academic Press, 662 pp.

LeBlond, P. H. and L. A. Mysak, 1978: *Waves in the Ocean.* Elsevier Oceanography Series, 602 pp.

Lindzen, R. D., 1967. Planetary waves on the beta-plane. *Month. Wea, Rev.,* **95(7),** 441-451.

Osychny, V. and P. Cornillon, 2004. Properties of Rossby waves in the North Atlantic estimated from satellite data. *J. Phys. Oceanogr.,* **31(1),** 61-76.

Pedlosky, J., 1979: *Geophysical Fluid Dynamics.* Springer-Verlag, 624 pp.

Rossby, C. G., 1939: Relation between variations in the intensity of the zonal circulation of the atmosphere and the displacements of semi-permanent centers of action. *J. Mar. Res.,* **2,** 38-55.

Figure caption

Fig. 1. The potential in the Schrödinger equation (2.9) – $(1+bz)^2$ – on the $-1 \le z \le 1$ interval for the indicated values of b. For $0<b<1$ the potential has no symmetry in this z-interval so that analytical solutions of the problem do not exist.

Fig. 2. The $E_0(\varepsilon, b)$ contours of the first eigenvalue of the Schrödinger equation (2.9). Along the b=0 and b=1 ordinates the numerically calculated values are exactly those anticipated analytically in Eqs. (4.1, for n=0) and (4.9), respectively: $E0(\varepsilon, 0)=1+(\varepsilon\pi/2)2$; $E0(\varepsilon, 1)=3\varepsilon$.

Fig. 3. The variation of E0(b) in the interval $0 \le b \le 1$ for the indicated (constant) values of ε. The upper panel describes the small ε regime, where E0 decreases monotonically with b and the lower panel describes the large ε regime, where E0 increases monotonically with b. In the narrow intermediate regime near ε=0.7 the slight variation of E0(b) is not monotonic throughout the entire $0 \le b \le 1$ interval.

Fig. 4. The dispersion relation, $\sigma(k)=kC(k)$, of Rossby waves in the new, b>0 theory and in the classical b=0 theory for the indicated values of $\delta\phi$ (=half the channel width divided by R) and α $(=gH/(2\Omega R)^2)$ and ϕ_0. The top panel compares the results of the two theories for a wide midlatitude channel while the bottom panel does the same for a realistic channel in the North Pacific spanning the latitudinal range of 11.5° – 51.5°.

Fig. 5. The eigenfunctions in the large ε case: ε=2; b=0.1 and $E_0=10.871$. The V(y) field in the upper left tile is common to the Rossby and (positive) Poincaré waves, whose u(y) and η(y) are shown in the upper right and lower left panels. The linear exponent of the Kelvin wave in the f-plane theory and the quadratic exponent of present theory differ only slightly as a result of the identical $\eta(y=-\delta\phi)=1$ normalization. The amplitude of the Kelvin wave decays with distance from the south wall, $y=-\delta\phi$, at a rate given by the radius of deformation $\alpha^{1/2}/\sin(1)=0.31$ so that at the north wall (located $2\delta\phi=0.312$ from the south wall) the amplitude of the Kelvin wave is e^{-1}.

Fig. 6. As in Fig. 5 but for the small ε case: ε=0.055; b=0.15 and E_0=0.862. These results should be compared with the classical, b=0, theory that yield E_0 = 1.00187 and purely sinusoidal V(y), u(y) and η(y) eigenfunctions. In new theory E0 is less than 1.0 and the eigenfunctions are very poor approximated by a simple sinusoidal variation.

Fig. 7. The divergence (δ/ik) and vorticity (ζ) fields for Rossby wave (upper panel) and (positive) Poincaré wave (lower panel) in the large ε case of Fig. 5. The inset in the upper panel shows the divergence at higher resolution from which it is clear that δ(y) is approximated by V(y). C_{rossby}=-0.00478; $C_{poincaré}$=2.789.

Fig. 8. The divergence (δ/ik) and vorticity (ζ) fields for Rossby wave (upper panel) and (positive) Poincaré wave (lower panel) in the small ε case of Fig. 6. The inset in the upper panel shows the divergence at higher resolution from which it is clear that δ(y) is approximated by V(y). C_{rossby}=-0.000103; $C_{poincaré}$=0.781.

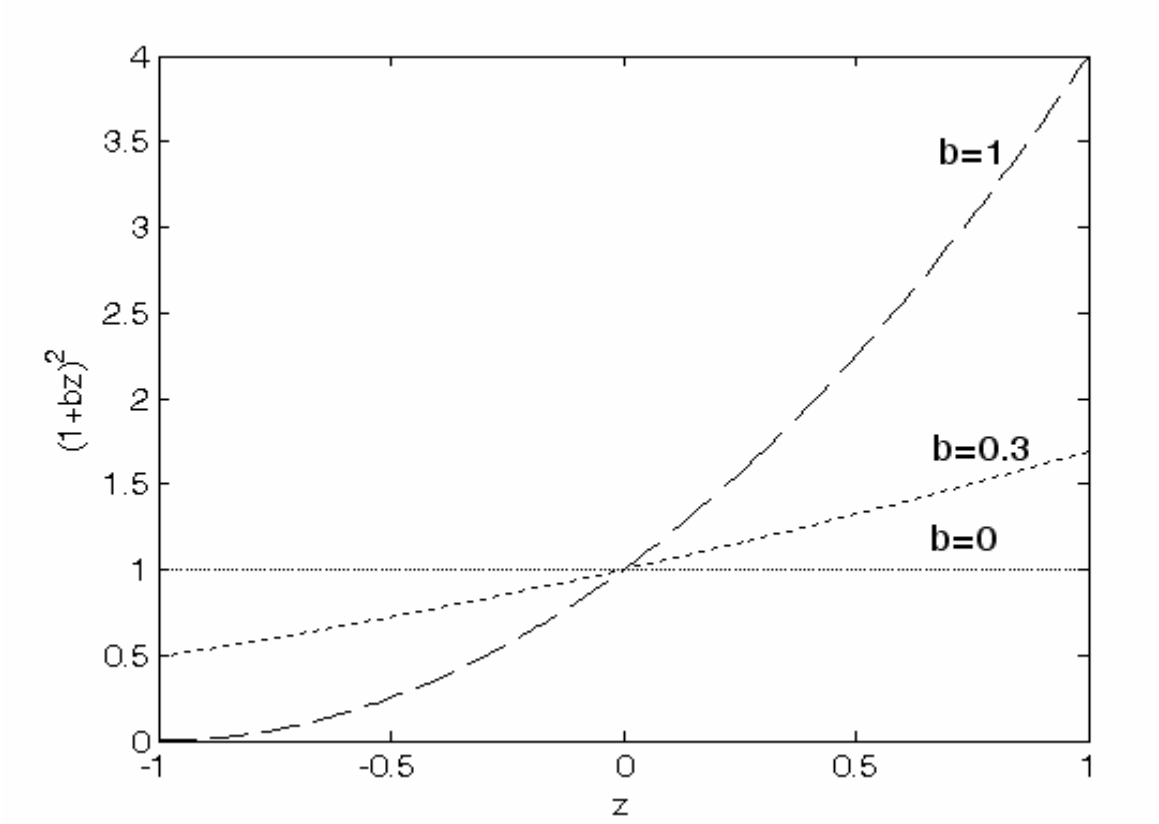

Fig. 1. The potential in the Schrödinger equation (2.9) – $(1+bz)^2$ – on the $-1 \leq z \leq 1$ interval for the indicated values of b. For $0<b<1$ the potential has no symmetry in this z-interval so that analytical solutions of the problem do not exist.

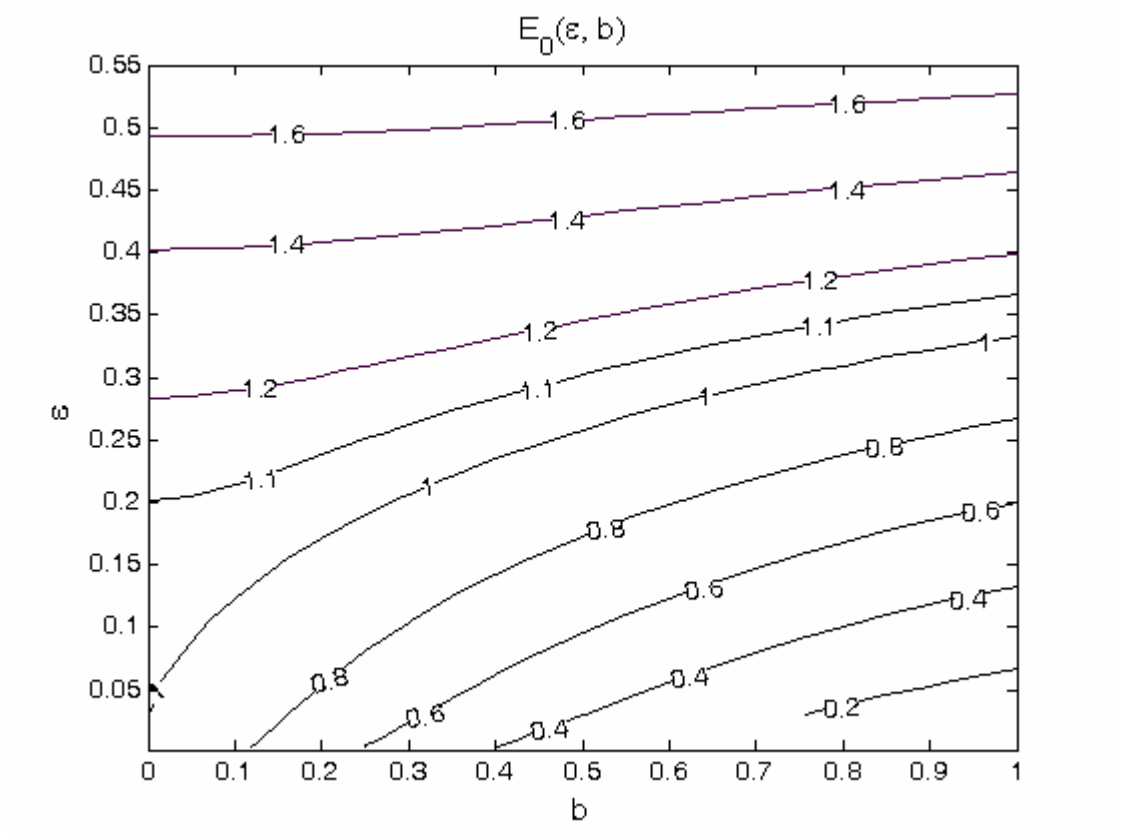

Fig. 2. The $E_0(\varepsilon, b)$ contours of the first eigenvalue of the Schrödinger equation (2.9). Along the b=0 and b=1 ordinates the numerically calculated values are exactly those anticipated analytically in Eqs. (4.1, for n=0) and (4.9), respectively: $E_0(\varepsilon, 0)=1+(\varepsilon\pi/2)^2$; $E_0(\varepsilon, 1)=3\varepsilon$.

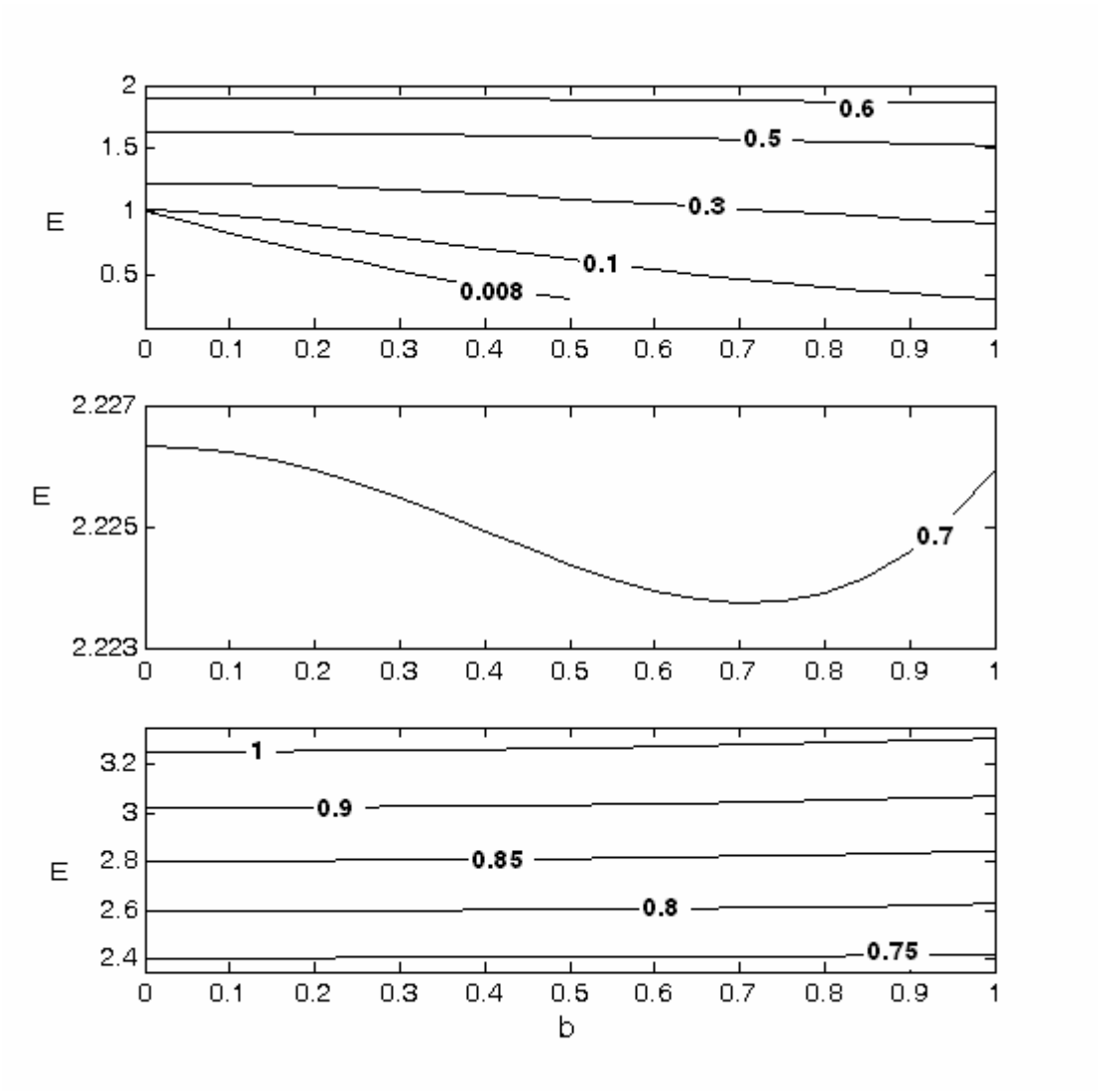

Fig. 3. The variation of $E_0(b)$ in the interval $0 \le b \le 1$ for the indicated (constant) values of ε. The upper panel describes the small ε regime, where E_0 decreases monotonically with b and the lower panel describes the large ε regime, where E_0 increases monotonically with b. In the narrow intermediate regime near $\varepsilon = 0.7$ the slight variation of $E_0(b)$ is not monotonic throughout the entire $0 \le b \le 1$ interval.

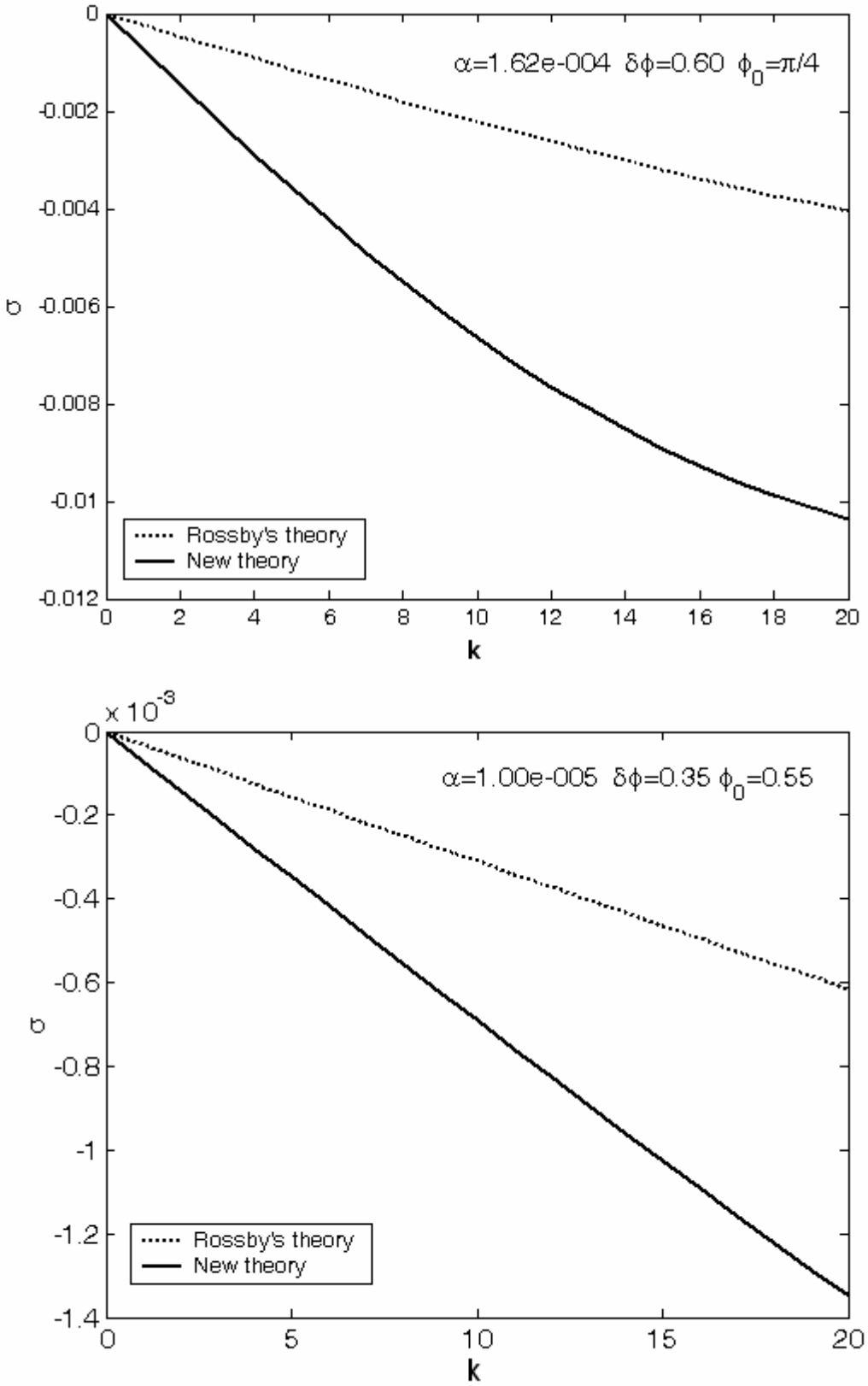

Fig. 4. The dispersion relation, $\sigma(k)=kC(k)$, of Rossby waves in the new, b>0 theory and in the classical b=0 theory for the indicated values of $\delta\phi$ (=half the channel width divided by R) and α (=$gH/(2\Omega R)^2$) and ϕ_0. The top panel compares the results of the two theories for a wide midlatitude channel while the bottom panel does the same for a realistic channel in the North Pacific spanning the latitudinal range of 11.5° – 51.5°.

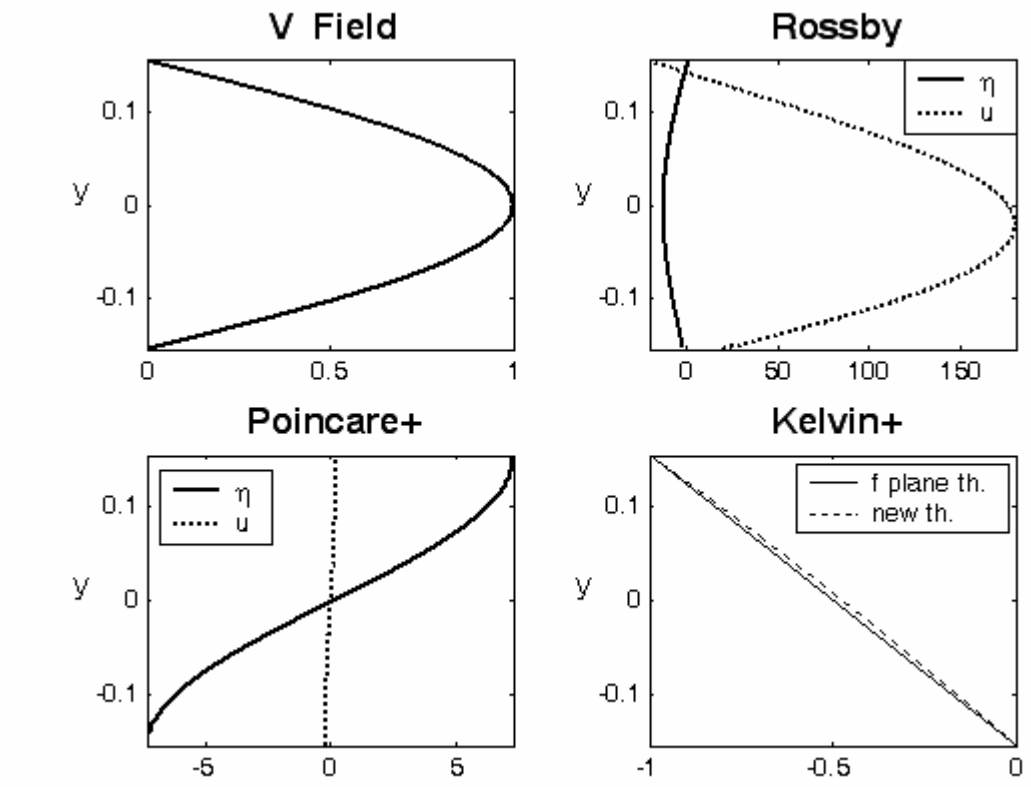

Fig 5. The eigenfunctions in the large ε case: ε=2; b=0.1 and E_0=10.871. The V(y) field in the upper left tile is common to the Rossby and (positive) Poincaré waves, whose u(y) and η(y) are shown in the upper right and lower left panels. The linear exponent of the Kelvin wave in the f-plane theory and the quadratic exponent of present theory differ only slightly as a result of the identical η(y=-δϕ)=1 normalization. The amplitude of the Kelvin wave decays with distance from the south wall, y=-δϕ, at a rate given by the radius of deformation $\alpha^{1/2}$/sin(1)=0.31 so that at the north wall (located 2δϕ=0.312 from the south wall) the amplitude of the Kelvin wave is e^{-1}.

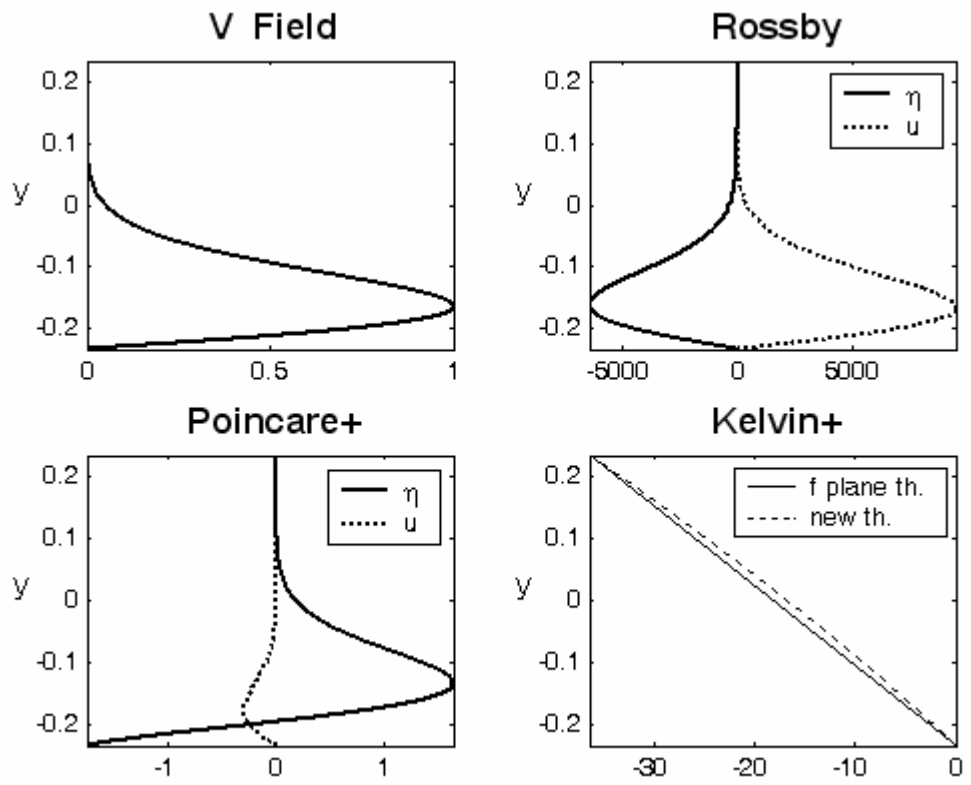

Fig. 6. As in Fig. 5 but for the small ε case: ε=0.055; b=0.15 and E_0=0.862. These results should be compared with the classical, b=0, theory that yield E_0 = 1.00187 and purely sinusoidal V(y), u(y) and η(y) eigenfunctions. In new theory E0 is less than 1.0 and the eigenfunctions are very poor approximated by a simple sinusoidal variation.

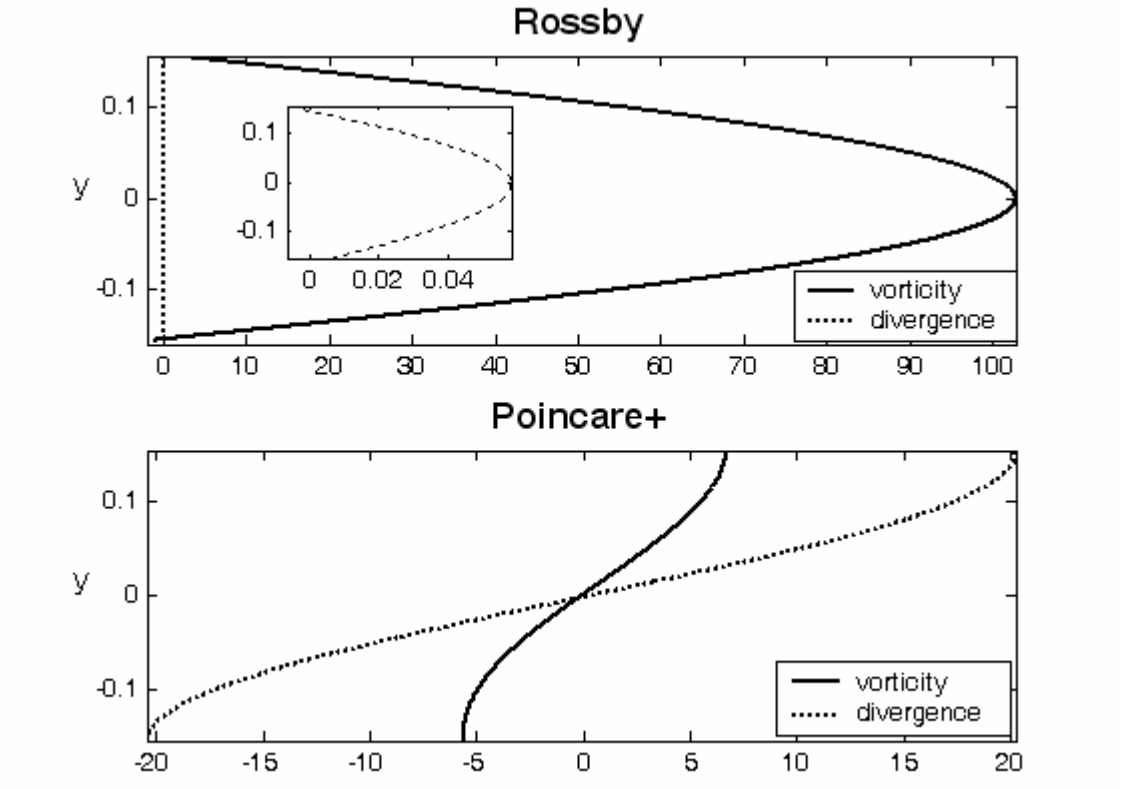

Fig. 7. The divergence (δ/ik) and vorticity (ζ) fields for Rossby wave (upper panel) and (positive) Poincaré wave (lower panel) in the large ε case of Fig. 5. The inset in the upper panel shows the divergence at higher resolution from which it is clear that δ(y) is approximated by V(y). C_{rossby}=-0.00478; $C_{poincaré}$=2.789.

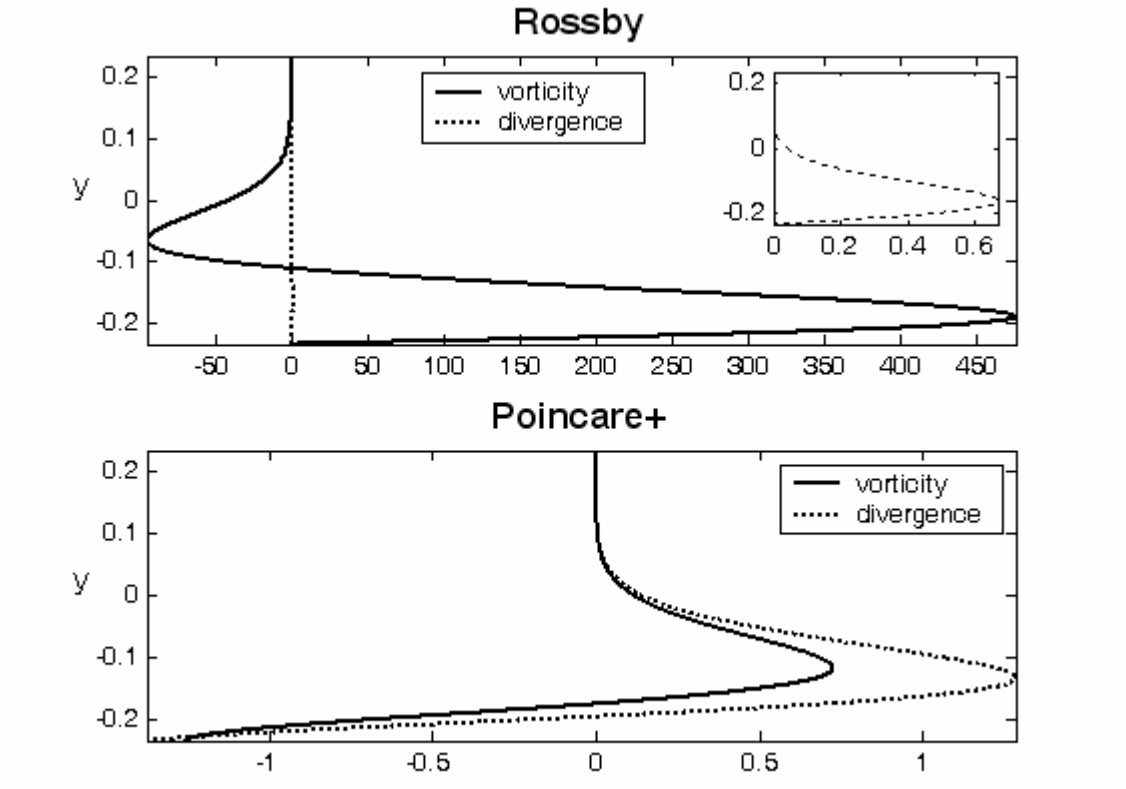

Fig. 8. The divergence (δ/ik) and vorticity (ζ) fields for Rossby wave (upper panel) and (positive) Poincaré wave (lower panel) in the small ε case of Fig. 6. The inset in the upper panel shows the divergence at higher resolution from which it is clear that δ(y) is approximated by V(y). C_{rossby}=-0.000103; $C_{poincaré}$=0.781.